Author's response to reviews

Title: Evaluation of dose reduction vs. standard dosing for maintenance of remission in patients with spondyloarthritis and clinical remission with anti-TNF (REDES-TNF): study protocol for a randomized controlled trial

Authors:

Caridad Pontes (cpontes@tauli.cat)
Jordi Gratacós (jgratacosmas@gmail.com)
Ferran Torres (ferran.torres@uab.es)
Cristina Avendaño (cavendano@salud.madrid.org)
Jesús Sanz (jesussanzsanz4@gmail.com)
Antoni Vallano (avallano@bellvitgehospital.cat)
Xavier Juanola (x.juanola@bellvitgehospital.cat)
Eugenio de Miguel (eugenio.demiguel@gmail.com)
Raimon Sanmartí (SANMARTI@ clinic.ub.es)
Gonzalo Calvo (gcalvo@ clinic.ub.es)

Version: 3
Date: 25 May 2015

Author's response to reviews: see over
Dear Sirs/Madams,

Please find attached below the answer to the reviewer and editor comments to the manuscript with reference MS:6429041331585359, entitled "Evaluation of dose reduction vs. standard dosing for maintenance of remission in patients with spondyloarthritis and clinical remission with anti-TNF (REDES-TNF): study protocol for a randomized controlled trial", whose authors are Caridad Pontes, Jordi Gratacós, Ferran Torres, Cristina Avendaño, Jesús Sanz, Antoni Vallano, Xavier Juanola, Eugenio de Miguel, Raimon Sanmartí and Gonzalo Calvo, and for which I am the corresponding author.

On behalf of the author’s please find a new version of the manuscript edited accordingly, as well as a point-by-point response to the concerns raised by the reviewer, which we feel have been all addressed. We would very much appreciate if you could re-consider the publication of this new version of the manuscript.

Please don't hesitate to contact me for any questions or clarifications.

With best wishes,

Jordi Gratacós, MD, PhD
Servei de Reumatologia, Hospital de Sabadell.
Institut Universitari Parc Taulí, Universitat Autònoma de Barcelona.
c/Taulí nº1, 08208 Sabadell (Barcelona), Spain.
jgratacosmas@gmail.com
Summary of major changes:
A new sentence has been included in page 9 (starting “If the remission rate of dose reduction is lower than ...”) and a new paragraph has been included in page 13 of manuscript (starting “The study has a number of potential limitations...”)”. All other edits have been made to the text and marked in the tracked version of the manuscript.

Reviewer’s comments
Major compulsory revision:
This is an important paper which should be published once the non-inferiority (NI) design is better defined and described. This would be aided by a biostatistical reviewer if not already done.

A NI design requires, for credibility, assurance that the active control would have beat placebo, had a placebo arm been present. The manuscript notes that it is anticipated that no less than 87% of patients allocated to full-treatment will show clinical remission after one year. What is needed is a body of past studies in support of this 87% figure. The only study quoted is ref 30, a very small cohort. Short of this body of studies, justifying a NI design is difficult, as success can mean that both the test and active control were efficacious, or it can mean that the test and active control were both inefficacious.

My concern is that, as written, the NI claim could be made even if the active control result was, say, 60% or even 40%. Would one make a NI claim if the full treatment result was 40% and the dose reduction result was, say, 30%? One way out of this dilemma was used recently in the rituximab trial in ANCA-associated vasculitis trial (RAVE), a NI setting where there was virtually no literature on the effect size of the active control, cyclophosphamide. In that trial a claim of NI when the test result was numerically lower than the active control result could only be made if the active control result itself was above a certain threshold (40% in the case of the RAVE study). This design was pre-agreed upon and it provided a defence against a NI claim when, in fact, it would be likely that neither arm was efficacious (see U Specks, et al. The Open Arthritis Journal 2011.4:1-18)

Author’s response
The reviewer points out the scarcity of data regarding the proportion of patients who, after achieving stable clinical remission, are expected to maintain remission for one year of full-dose treatment with antiTNF. In addition, the reviewer is concerned about how this uncertainty might hypothetically allow non-inferiority to be attributed to an
inefficacious treatment, unless a lower boundary is established for the efficacy of full dose treatment, similar to the methodology previously applied in similar trials (U Specks, et al. The Open Arthritis Journal 2011.4:1-18).

We feel this is a different situation: the RAVE trial compared two drugs for both of which uncertainty on actual efficacy was recognised, and the REDES trial will compare different maintenance dosages of the same product to which the patient has previously responded during induction of remission; the current recommendation of use of the product is to keep long term maintenance with full dose. Nevertheless, we acknowledge that our estimate of the response to the maintenance treatment with the full dose is based only on a small cohort.

In order to address any concern on lack of efficacy, we will introduce, in the statistical analysis plan, the following requirements to conclude non-inferiority of dose reduction:

- First, the lower boundary of the two-sided 95% confidence interval of the difference between dose reduction vs. full dose patients achieving the primary outcome will have to be greater than the -17% non-inferiority margin.

- Second, if the remission rate of dose reduction is lower than that of the full dose, the lower boundary of the confidence interval in the full dose arm will have to be > 60%, to ensure that the control treatment has been reasonably effective.

**We now state:**

“The principal and key secondary end-points will be assessed by estimating the between-treatment risk differences after 1 year of randomization and checking these against the pre-defined non-inferiority margin (delta (δ)) of 17%. If the remission rate of dose reduction is lower than that of full dose, the lower bound of the confidence interval in the full dose arm has to be above 60% to conclude non-inferiority, to ensure that the control treatment has been reasonably effective. Rates and risk differences will be estimated using a log-binomial regression model including the treatment and the factor used to stratify the assignment. In the event that the model does not fit, the Poisson link distribution function with robust variance will be used instead [34-38].
Additional questions/comments:

1-In Study Objectives (p5) you say “patients with axial non-psoriatic spondyloarthritis” yet psoriasis is not an exclusion in Table 2. Please clarify.

Author’s response

The exclusion criteria indicate that secondary spondyloarthritis, which also refers to psoriatic arthritis with axial involvement, is an exclusion criterion. “non-psoriatic” has been deleted to avoid confusion.

We now state:

“The primary study objective is to assess, in patients with axial spondyloarthritis who have achieved sustained clinical remission with antiTNF, whether the proportion of patients who, after one year reach an acceptable therapeutic goal (BASDAI <4, physician global assessment <4 and patient <4 and axial night pain <4) [27] is not inferior in patients receiving reduced doses of antiTNF than in patients using standard antiTNF doses, according to the summary of product characteristics. “

2-Study design (p5) entry requires anti-TNF treatment for a minimum of 12 weeks then in remission for an additional 8 weeks, so 20 weeks treatment is needed. I assume all enrollees are still on anti-TNF therapy, yes? In other words, no patient has already been withdrawn from their TNF? Is this correct?

Author’s response

The requirement is that the patient is in stable remission for a minimum of 8 weeks after completing the induction phase of treatment (lasting 12 weeks); thus 20 weeks is the minimum exposure time to be eligible.

All patients are on antiTNF treatment for the whole study duration – the only difference between groups will be the dose given (full or reduced).

We have re-written the text (in blue) to improve clarity on this aspect.

We now state:

“A prospective, multicenter, controlled, randomized open-label study was designed. Patients with axial spondyloarthritis treated with antiTNF for ≥ 20 weeks with sustained clinical remission during the last ≥ 8 weeks will be included.
3-The proviso for three populations is confusing. One usually has the all randomized population (ITT population), and the per protocol population, and with a NI trial both are important as is noted. If the difference between the RA and the FAS amounts to missing data, then you need a sensitivity analysis to demonstrate that the missing data are not informative.

Author’s response

We have prospectively detailed in the statistical analysis plan (SAP) the valid causes for excluding patients from the RS to define the FAS and, in essence, we can argue that all of these reasons are fully compliant with the ICHE9 (CPMP/ICH/363/96); please find below the specifications given in the SAP regarding the definition of the populations.

We do not expect many drop-outs between the RS and FAS populations. However, in addition to the predefined analyses of the PP and FAS and following the reviewer’s suggestions, we will also test the RS by imputing to failure any dropouts without valid measurements for the primary outcome.

1. **Randomized population:** All study patient data will be analyzed as randomized.

2. **Full Analysis Set (FAS):** All patients who are randomized and who have initiated the study medication will be included in the FAS population. The accepted exclusions, as per the ICHE9, are predefined as follows:

   - Subjects who fail to satisfy an entry criterion may be excluded from the analysis without the possibility of introducing bias only under the following circumstances:
     
     i) the entry criterion was measured prior to randomization;
     
     ii) the detection of the relevant eligibility violations can be made completely objectively;
     
     iii) all subjects receive equal scrutiny for eligibility violations; (This may be difficult to ensure in an open-label study, or even in a double-blind study if the data are unblinded prior to this scrutiny, emphasizing the importance of the blind review.)
     
     iv) all detected violations of the particular entry criterion are excluded.

   - **Per Protocol Population:** Per protocol (PP) patient sets will be defined as patients included in the FAS set who took the study medication without major
protocol deviations that might impact the study’s main assessments. These deviations will be assessed during the data blind review prior to database lock.

References


We now state:

“The PPS is the pre-defined primary population for this non-inferiority study. However, the principal end-point and the key secondary end-points will also be tested in the RS and FAS populations for consistency.

(...) Therefore, although analysis of the PP set will be the predefined primary analysis, the principal end-point and the key secondary end-points will also be tested in the RS and FAS populations. “

4-You note that “in the event the model does not fit”. Please explain by what criterion this is to be made. You then say the Poisson link distribution function...will be used. Explain why.

Author’s response

It is known that the log-binomial model is less numerically stable than the logistic model and there are times when the log-binomial model fails to converge. The model might fail to converge either because of bad initial values, for example, initial values not in the restricted parameter space, which can be easily fixed by specifying better initial values, or because the maximum likelihood solution occurs on the boundary of the parameter space. In the latter situation, the derivative of the likelihood at its maximum may not be 0. Thus, standard software packages which maximize the likelihood by finding the point at which the derivative is equal to 0 (namely Newton’s method) may not find the solution (Deddens 2008; Petersen 2008). When this is the case, use of the Poisson regression capability with robust variance has been recommended (Skov 1998; Zou 2004, Spiegelman 2005, Bieler 2010). One can think of a sample of binomial data (0 or 1) as being approximately Poisson, where the probability of a value of 2 or greater is low enough that no values greater than 1 occurred in the sample obtained. If the logarithm of the Poisson parameter is assumed
to be linearly related to a set of independent variables, then the exponentiation of any coefficient of the model will yield an estimate of a ratio of Poisson parameters, adjusted for the other covariates. Because the observed data consist of only 0s and 1s, this ratio can be used as an approximation to the adjusted rate. Poisson regression has standard errors which are too large and while the estimates are valid, they are not as efficient when compared with these log-binomial maximum likelihood estimators (Deddens 2008).

With respect to our trial, in principle we do not expect that lack of convergence will be a problem because of: (a) the trial size (n≈170), (b) the small number of variables included in the model (treatment and stratum -i.e. previous treatment-) and (c) although convergence problems can occur with qualitative data, they are much more common with quantitative data (Petersen 2008). However, we decided to predefine an alternative to give more robustness to the statistical plan and consequently to the trial results. In the hypothetical case of a lack of model convergence, this would be assessed by the inspection of the log information provided by the SAS system, in addition to not being able to obtain the model estimates (plus standard errors). In this a priori unexpected scenario, we would then proceed with the alternative model which, in addition, is more conservative for the study purposes.

In essence, we think that the primary model and the strategy for covering a hypothetical lack of model convergence are reasonable and in line with the literature recommendations.

References

- Petersen MR, Deddens JA. A comparison of two methods for estimating


5-There should be a “study has weaknesses” paragraph in the discussion.
6-In the discussion: Query the use of blinded assessors.
7-In the discussion: Discuss the underpowering for inferences about a specific TNF.

Author’s response

We now state:

“The study has a number of potential limitations. First, the study is not blinded, since it was thought that the complexity, costing and risk of medication errors associated with the use of double dummy (different doses in iv injections for infliximab, placebo subcutaneous injections matching the alternative posology for adalimumab, etanercept and golimumab) did not outweigh the benefits of blinded assessment of outcomes, especially for the subcutaneous treatments, which would have required an increase in the number of injections over a long time period. Likewise, blinded assessors will not be used because the decision to modify treatment in the event of clinical flare will mainly be based on the patient’s reporting of disease signs and symptoms (BASDAI score and axial night pain), and it is expected that if the patient’s reports are biased due to knowing the identity of treatment, they will tend to consider the low-dose treatment as less efficacious than the full dose treatment. Thus, the bias would likely be conservative, contrary to the main study hypothesis, and reflective of what may be expected after dose reduction in routine clinical practice. Secondly, the pragmatic approach means the use of NSAIDs and DMARDs will not be standardized, and may be varied by the investigators as required during the study to control symptoms. While this may be regarded as a potential source of confusion, the use of these additional treatments is quantified during the trial, and will be analyzed to determine whether either of the arms is associated with an increase in the use of concurrent anti-inflammatory drugs or DMARDs. Thirdly, patients will be stratified only according to the antiTNF drug, in order to maintain a reasonable number of strata. Additional factors with potential prognostic implications (e.g. time since first diagnosis, use of DMARDs, duration of clinical remission at randomization) were not considered. Although these factors have been reported to predict a clinical response
to antiTNF therapy [55,56], until now there is no data on their predictive value in predicting the clinical result of antiTNF dose reduction. Fourthly, the duration of follow-up until the main efficacy assessment is limited to one year. Although this period may be considered a reasonable timeframe in which to test the clinical acceptability of each therapeutic option, the study will not be able to provide information to guide clinical decisions after this time, and it is unlikely the study will be able to detect relevant differences in structural end-points by spine imaging, since changes are generally slower in axial spondyloarthritis. It is also anticipated that the study will not be able to detect differences in efficacy or safety between the dosages of each antiTNF studied.”

New references:


Stylistic suggestions

1-Avoid self-aggrandizement: E.g., in the abstract/discussion: The REDES-TNF study is a well designed and pragmatic .... I’d delete “well-designed”. Similarly throughout the manuscript.

Author’s response

The text has been reviewed to eliminate self-engrandizement. All changes are marked in the tracked version of the re-submitted version.

2-Edit the entire manuscript by someone with extensive English experience. For example, The first sentence in the last paragraph is better written as: In summary, the REDES-TNF ... aimed to answer the real need for evidence to support medical decisions now taken empirically

Author’s response

The manuscript has been reviewed by a native English medical translator. All changes are marked in the tracked version of the re-submitted version.
**Editor’s comments**

1. Please note that XV should be XJ in the Authors' Contributions section.

**Author’s response**

The text has been corrected. All changes are marked in the tracked version of the re-submitted version.

2. For additional files, please ensure that you list the following information after your reference section in your manuscript:

**Author’s response**

The text has been corrected to refer to additional file 1 instead of supplementary table 1. All changes are marked in the tracked version of the re-submitted version.

**We now state:**

**ADDITIONAL FILES:**

File name: Additional file 1. doc

Title: Supplementary table 1

Description of data: Listing of the Ethic’s Committees that reviewed and approved the study protocol.

File name: Additional file 2.doc

Title: SPIRIT_Fillable-checklist

Description of data: SPIRIT_Fillable-checklist completed for the submitted protocol